

Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be?

Year: 2023 | Citations: 181 | Authors: I. Alberts, L. Mercolli, T. Pyka

Abstract

There has been substantial press recently regarding the impressive performance of large language models (LLM), particularly the OpenAI tool “Chat Generative Pre-Trained Transformer,” commonly known as “ChatGPT” [1]. LLM represent artificial intelligence (AI) tools based on multilayer recurrent neural networks that are trained on vast amounts of data to generate human-like text (<https://ai.googleblog.com/2017/08/transformernovelneuralnetwork.html>). Whereas traditional language models are programmed to use statistical techniques to predict the next word in a sentence, ChatGPT uses transformer-based models that allow for the processing of vast amounts of data in parallel. The result is a revolution in the ability of these models to understand and generate text. Their performance is remarkable, e.g., ChatGPT is capable of writing lines of code, producing plays, stories, poetry as well as simulated scientific content such as abstracts. While there has been much fanfare in the media regarding this undoubtedly impressive performance, there is much less information available about how this might affect the nuclear medicine community, or its reliability in producing nuclear medicine and molecular imaging-related content. It is currently unclear to what extent ChatGPT might help as a collaborative tool, for example correcting or helping to improve upon a researcher’s writing or as a tool for summarising nuclear medicine literature. Within seconds, ChatGPT is capable of producing convincing and grammatically fluent prose that is indistinguishable from content produced by human researchers. The threat that this poses to academic publishing models is already apparent [2]. Controversially, ChatGPT has (at the time of writing) already been listed as a co-author on four academic publications [3]. Anecdotally, students are already using the tool as a writing assistant, raising issues of academic integrity and plagiarism [4]. There are already 25 PubMed entries for “ChatGPT”, this will likely grow rapidly in the coming weeks and months. In response, a number of journals are already implementing editorial policies about the acceptability of AI-assisted writing or clarifying issues around authorship [3, 5]. Some internet fora have already banned ChatGPT-generated answers owing to their unreliability (<https://meta.stackoverflow.com/questions/421831/temporarypolicychatgptisbanned>). Recent experience has shown that AI tools can be harnessed to mass-produce questionable content on social media networks or social media bots that can deliberately amplify misinformation [6]. This experience might portend the future of ChatGPT-generated academic content. A report from the Copenhagen-based Institute for Future Studies estimates that 99% of the internet could be produced by generative AI by 2025 (<https://cifs.dk/news/what-if-99-of-the-internet-was-made-by-ai>). At present, ChatGPT is not capable of producing an entire research paper sua sponte, although it is predicted and indeed conceivable that this might soon be the case [7]. Nevertheless, it can already, even in the currently available beta version, produce a very convincing abstract [8]. We wonder whether conferences might soon be flooded with AI-generated abstracts or whether predatory publishers [9] might be catalysed by the ability of ChatGPT to churn out convincing but ultimately unreliable content. Even the review process could be influenced: there are already proposals to harness the ability of LLM to summarise text as a tool for the sifting out of low-quality studies submitted to a journal. Once can imagine a not-too-distant future where AI might generate and review research [10]. This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence)